Attorney Docket No.: PHO 99004CIP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: *Bennett*	)	Art Unit: *2626*
Serial No.: *10/653,039*	)	Examiner: *Martin Lerner*
Filed: August 29, 2003 as continuation-in-part of 09/439,145 filed November 12, 1999, now U.S. patent 6,633,846	)	
For: *Query engine for processing voice based queries including semantic decoding*	)	

Appeal Brief filed under 37 C.F.R. § 1.192

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Appellant submits this Appeal Brief under 37 C.F.R. § 41.37.

A short introduction of the prosecution history is first presented. This brief also contains the following sections as required by 37 C.F.R. § 41.37 and MPEP § 1206:

I. Real Party In Interest

II. Related Appeals and Interferences

III. Status of Claims

IV. Status of Amendments

V. Summary of Claimed Subject Matter

VI. Grounds of Rejection to be Reviewed on Appeal

VII. Grouping of Claims

VIII. Argument

IX. Claims

X. Evidence

XI. Related Proceedings

Appendix A Claims

BRIEF INTRODUCTION AND REVIEW OF PROSECUTION HISTORY

The pending claims include claims 1-15 and 22 – 28. The present application is a continuation-in-part of application serial no. 09/439,145 filed November 12, 1999, now U.S. Patent No. 6,633,846.

The claims were rejected in a First Office Action dated March 2, 2007. In this action the Examiner rejected claims 1-3, 5, 9-10, 22 and 24-25 under § 103 based on the Junqua et al. reference (U.S. Patent No. 6,314,398) taken with Philips et al. (6,519,562). Claims 6, 11 – 13 and 27 – 28 were rejected in light of such references taken with Barclay et al. (U.S. Patent No. 5,960, 399) while claims 7, 8, 14, 23 and 26 were rejected in light of such references taken with Appelt et al. (6,601,026), claim 15 was rejected in light of such references taken with Agarwal et al. (5,842,196), and finally claim 4 was rejected in light of such references taken with McDonough et al. (5,625,748).

In Amendment A filed August 14, 2007 the Applicant explained that the primary reference (Junqua et al.) did not in fact show search predicates, but only key word searching. The Applicant further amended the independent claims (1, 22) to clarify that the "search predicates" of the present claims corresponded to logical operators (i.e, as an example, a NEAR operator, an AND operator, etc.). The Applicant further pointed out that the secondary reference relied upon (Philips et. al.) was relevant only for purposes of teaching statistical processing for speech recognition and *not* natural language processing.

The Examiner repeated the rejection in a second Office Action, but this time relied on a newly cited reference – Kupiec (U.S. Patent No. 5,500,920) as the primary reference taken with the aforementioned Junqua et al. reference. In addition, another new reference – Joost was applied to claims 10 – 12 and 27 – 28. The Examiner made the action "Final" based on his conclusion that the amendment to independent claims 1 and 22 had raised "new issues."

In an Amendment B filed after Final on November 28, 2007 Applicant challenged the propriety of the Finality of the second Office Action. As Applicant explained, if the newly cited Kupiec reference indeed disclosed all elements of the amended independent claims 1, 22 then by fundamental principles of patent law it should have

been cited by the Examiner earlier as anticipating the original broader versions of such claims. Instead Applicant did not have an adequate opportunity to address the reference at the appropriate time. The Applicant again amended the claims to note that the natural language engine was able to generate two different types of logical operators, and again argued other distinctions over the references.

The Examiner issued an Advisory Action on December 20, 2007 in which he entered the Amendment but maintained the rejections based on Kupiec taken with Junqua et al. again. The Examiner further maintains that the citation of Kupiec in the Final Office Action was necessitated by amendment.

I. REAL PARTY IN INTEREST

Phoenix Solutions, Inc. is the assignee of the present application and the real party in interest.

II. RELATED APPEALS AND INTERFERENCES

There are no other appeals, interferences or judicial proceedings known to Appellant, Appellant's legal representative, or the Assignee of the present application which will directly affect or be directly affected by or have a bearing on the Board's decision in this appeal.

III. STATUS OF CLAIMS

Claims 1 – 15 and 22 - 28 are pending. Claims 1 and 22 are independent. A complete copy of the pending claims is provided in Appendix A.

IV. STATUS OF AMENDMENTS

There is no amendment which has not been entered in this case. The claims are as submitted in Amendment B filed November 28, 2007 and entered by the Examiner on December 20, 2007.

V. SUMMARY OF CLAIMED SUBJECT MATTER

The citations below are solely presented to help communicate aspects of the support for limitations of the claims as required by the regulations for Appeal briefs and are not intended to be exhaustive or suggestive of the legal interpretation to be given to such claims.

A. Independent claim 1

Independent claim 1 recites a speech query recognition system comprising:

a speech recognition engine for generating recognized words taken from an articulated speech utterance; (see among other places, FIG. 1, boxes 155 and 182 (alone or together) and discussion for FIGs. 2A; FIG. 2B; FIG. 2C; FIG. 2D; FIG. 3; FIG. 4; FIG. 4A; and accompanying discussion at pages 18, ll. 8 – 18; page 20 – l.18 to page 26, l. 23; page 30, l. 30 to page 31, l. 10; page 33, l. 21 to page 36, l. 31; page 39, l. 19 to page 42, l. 12)

a natural language engine configured for linguistically processing said recognized words to generate at least two different types of search predicates for said articulated speech utterance; (see among other places, FIG. 1 box 190; and discussion for FIGs. 4D; FIG. 5 and FIGs. 11A, 11B and 11C at page 19, ll. 30 – 31; page 26, l. 24 – page 27, l. 17; page 29, ll. 2 – 19; page 37, ll.15 – 22; page 42, ll. 14 – 28; as well as FIGs. 19 – 21 and discussion at page 55, l.13 – page 62, l.11; page 64, ll. 16 - 25)

wherein said search predicates correspond to logical operators to be satisfied by a potential recognition match; (id.)

a query formulation engine adapted to convert said recognized words and said search predicates into a structured query suitable for locating a set of one or more corresponding recognized matches for said articulated speech utterance; and (see among other places FIG. 1, box 184; and discussion for FIG. 4B; FIG. 4C; FIG. 4D; FIG. 5; FIG. 10; and discussion at page 27, l. 20 – page 30, l. 6; page 42, l. 30 to page 47, l. 25)

said natural language engine further being configured for linguistically processing said set of one or more corresponding recognized matches to determine a final match for said articulated speech utterance using both semantic decoding and statistical based processing performed on said recognized words. (See FIGs. 19 – 21; and discussion at page 37, l. 27 – page 38, l. 29; page 47, l. 27 to page 51, l. 14;

and more particularly at pages 55 – 64 for discussion of semantic decoding and statistical based processing)

B. Independent claim 22

Independent claim 22 recites a method of recognizing a speech query comprising the steps of:

(a) recognizing text in an articulated speech utterance; (see claim 1) and

(b) linguistically processing said recognized text to generate at least two different types of search predicates for said articulate speech utterance; (see claim 1)

wherein said search predicates correspond to logical operators to be satisfied by a potential recognition match; (see claim 1)

(c) generating a query to identify a potential match for said speech utterance, said query being based on said recognized text and said search predicates; (see claim 1)

(d) determining a final match for said speech utterance by linguistically comparing any potential matches identified by said query with said articulated speech utterance; (see claim 1)

wherein both semantic decoding and statistical based processing operations are used to determine said final match. (see claim 1)

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The issues presented for appeal are:

- whether claims 1 – 3, 5, 9, 22 – 25 are obvious in light of Kupiec (5,500,920) taken with Junqua et al. (6,314,398) et al.;
- whether claim 4 is obvious in light of Kupiec (5,500,920) taken with Junqua et al. (6,314,398) et al. and McDonough et al. (5,625,748);
- whether claims 6 and 13 are obvious in light of Kupiec (5,500,920) taken with Junqua et al. (6,314,398) et al. and Barclay et al. (5,960,399);
- whether claims 7 – 8, 14 and 26 are obvious in light of Kupiec (5,500,920) taken with Junqua et al. (6,314,398) et al. and Appelt et al. (6,601,026);
- whether claims 10 – 12 and 27 – 28 are obvious in light of Kupiec (5,500,920) taken with Junqua et al. (6,314,398) et al. and Joost;
- whether claim 15 is obvious in light of Kupiec (5,500,920) take with Junqua et al. (6,314,398) et al. Agarwal et al. (5,842,196).

VII. GROUPING OF CLAIMS

The claims do not stand or fall together because they are directed to different facets of the present inventions and/or are more particularly directed to specific features of such inventions. A detailed discussion of such differences is given below.

VIII. ARGUMENT

All of the present claims are rejected under a single §103 rejection based on the combination of Kupiec (5,500,920) taken with one or more secondary references including the Junqua et al. reference.

As a preliminary matter, the Examiner acknowledges that Kupiec et al does not disclose a natural language engine as set out in claim 1. Instead, the Junqua et al. reference is cited to support the notion that one skilled in the art would incorporate the latter's teachings on natural language processing to address and rectify the deficiencies in Kupiec. This augmentation of Kupiec's disclosure, however, is not sufficient to render the claims of the present application obvious for the reasons set forth in more detail below.

In brief, the combination of these references fails for the following reasons. First, speech recognition (SR) processing and natural language (NL) processing are fundamentally different techniques with different goals. The goal of SR generally is merely to decipher the actual words spoken in an utterance. The goal of NL processing generally is to comprehend the meaning of the utterance, and this requires an entirely different set of designs, strategies and techniques. As set out in the disclosure:

> In contrast to word recognition, Natural language processing (NLP) is concerned with the parsing, understanding and indexing of transcribed utterances and larger linguistic units. Because spontaneous speech contains many surface phenomena such as disfluencies, - hesitations, repairs and restarts, discourse markers such as '*well*' and other elements which cannot be handled by the typical speech recognizer, it is the problem and the source of the large gap that separates speech recognition and natural language processing technologies. Except for silence between utterances, another problem is the absence of any marked punctuation available for segmenting the speech input into meaningful units such as utterances. See page 4, ll. 21 – 28.

This distinction helps establish why the references fail to disclose the claimed limitations, including the following:

1) Neither reference discloses "semantic decoding" as set out in the claims;
2) Junqua et al. discloses a form of natural language processing, but nothing which employs both "semantic decoding" and "statistical processing"; In Kupiec et al. any such related functions are limited to a speech recognition engine;
3) Junqua et al. discloses a form of natural language processing, but nothing which generates "search predicates", including "logical operators"; while Kupiec's speech recognition engine appears to disclose search predicates, he does not do so based on "recognized' words and there is no evidence to suggest that one skilled in the art would or could modify Kupiec to include a natural language engine with such qualities;
4) Junqua et al. discloses a form of natural language processing, but nothing which linguistically processes a "….set of one or more corresponding recognized matches…" Kupiec also does not linguistically process any matches; he simply retrieves them based on co-occurrence of certain words in the query.

The Examiner's evidence for obviousness is summarized in the following sentence at page 4:

> "…It would have been obvious to one having ordinary skill in the art to construct a search query of Kupiec ('920) with semantic constraints of a natural language engine as taught by Junqua et al. for a purpose of permitting a user to communicate in any style of language."

The Examiner's rationale, therefore, appears to rely strictly on element (C) of the Examination Guidelines for Determining Obviousness Under 35 U.S.C. 103 in View of the Supreme Court Decision in KSR International Co. v. Teleflex Inc. which states that one rationale for obviousness can be based on

> *"…(C) Use of known technique to improve similar devices (methods, or products) in the same way;"*

No other rationale is provided by the Examiner for obviousness, and therefore such other potential arguments are not addressed at this time.[1]

The problem with this rationale is that it does not find any support in the references and is inapplicable in this case. For example, as indicated above, the Examiner does not even contend that the NL engine in Junqua et al. has the requisite capabilities. He simply refers in the rejection to the "semantic constraints" taught by the latter; but a review of such disclosure certainly does not inform one skilled in the art that Junqua et al. is referring to "semantic decoding" of recognized words within the utterance as taught by the present claims.

The Examiner's rejection therefore relies on an argument that requires one skilled in the art to *incorporat*e and *modify* the incomplete NL processor from Junqua et al. to nonetheless perform an analogous set of functions performed by the SR processor in Kupiec. Applicant submits that there is no evidence that such combination would ever be reasonably contemplated by one skilled in the art.

Conversely it can be further seen that there is absolutely no guidance or information to instruct one skilled in the art on *how* they would duplicate the processing functions offered by the Kupiec speech recognition engine within the Junqua et al. natural language engine. This is not surprising, since, as noted earlier, the goals and designs of SR engines and NL engines are entirely different.

Thus, the implementation of certain features within an SR engine is not apt or appropriate in most cases for inclusion in a NL engine. This means that to a large extent the suggested modifications by the Examiner would not be compatible or achievable with the disparate systems disclosed in Kupiec et al. and Junqua et al. This incompatibility presents yet another reason for non-obviousness.

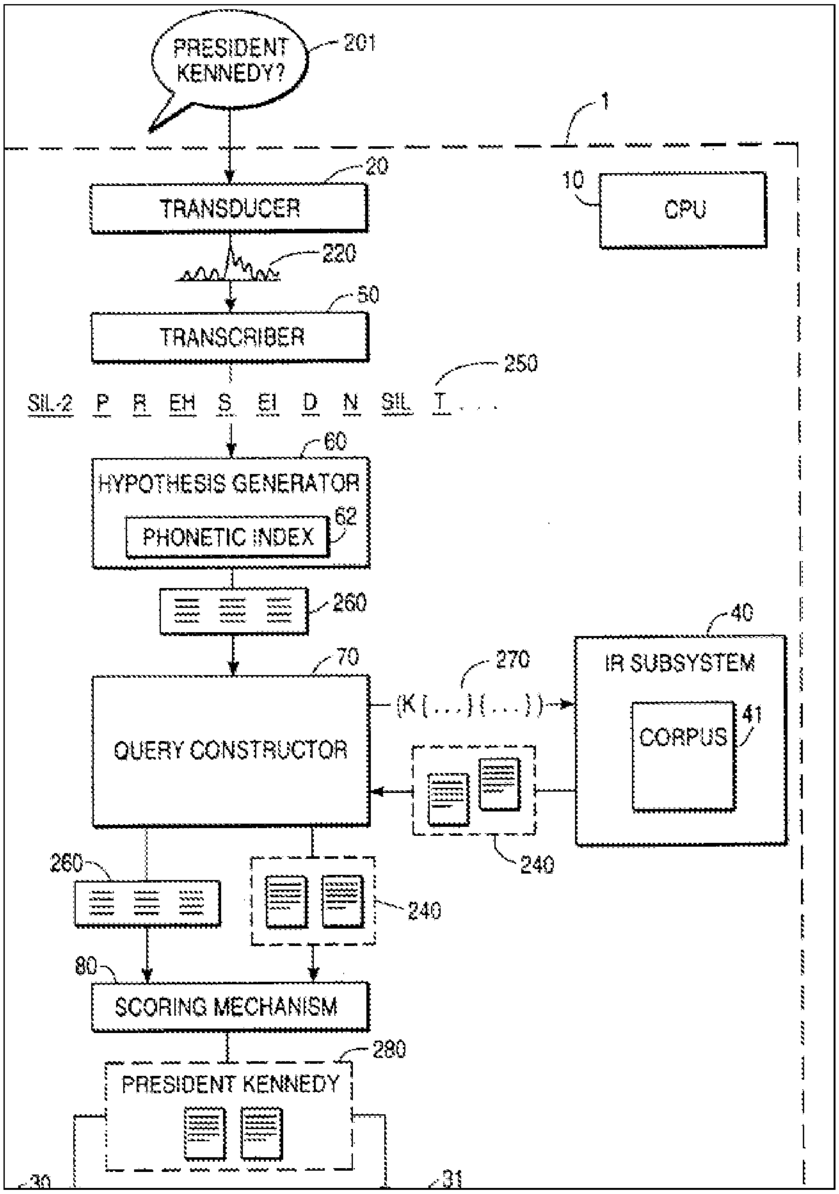
The references and arguments are now addressed in detail.

Claim 1 and the rejections based on the Kupiec & Junqua References

The Kupiec reference is directed to an improved speech recognition system, which basically tries to identify the exact words spoken by a user by comparing the utterance to text found in a corpus of documents. There is no discussion whatsoever of attempting to determine the meaning of the user's utterance. An excerpt of FIG. 1 of Kupiec is noted below:

(..continued)

[1] In other words, the Examiner has not presented any teaching, suggestion or motivation evidence, or any argument that the present case should be resolved under the "predictable results" rationale test set out in KSR.



Thus the hypothesis generator 60, the query constructor 70, the IR subsystem 40, the scoring mechanism 80, etc., are all part of a speech recognition engine. In other words, <u>Kupiec</u> is effectively a reference that is trying to increase the accuracy rate of a literal <u>transcription</u> of an utterance – i.e., the actual words that the user spoke. This can be confirmed from reviewing <u>Kupiec</u> at col. 2, ll. 3 – 12; and from other passages, including column 3:

> The invention uses information, particularly co-occurrence information, present in the corpus to help it *recognize what the user has said.*

-- Neither Kupiec et al. nor Junqua et al. show "semantic decoding"

Kupiec basically constructs hypotheses of what the user said, and then compares such hypotheses against text it finds in the corpus of documents. The reference relies on a strategy in which it is assumed that the proximity of two words to each other in an utterance is a useful indicator for determining what the user said, and such proximity is used to find corresponding pairs in the corpus. As explained in Kupiec:

> The invention assumes that semantically related words in the speaker's utterance will tend to appear together (co-occur) more frequently in the corpus.

Thus Kupiec states that in the utterance "President Kennedy" for example, the determination is made that the two words, President, and Kennedy, are somehow *semantically* related simply because they are within a certain proximity to each other in the utterance. This is the extent of the "semantic" processing disclosed by Kupiec, which can be seen is geared entirely to the operation of a speech recognition engine whose goal is to determine *what* the user said. In this respect therefore the definition used by Kupiec of what constitutes a semantic relationship – i.e., two proximate *words* – is again adapted and limited to such environments, and there is no indication that such technique would be useful or applicable to a natural language engine.

In contrast present claim 1 makes note that "semantic decoding" is done on the recognized words from the utterance. As explained in the specification, this decoding can be done in different ways, one of which is:

> ….[F]or example, if the stored question is: '*How do I reboot my system*', and the user's question is: '*What do I do when my computer crashes'*, we could, with the help of a lexical dictionary such as WordNet, establish that there is a semantic relationship between '*computer crash*' and '*rebooting*'. This would then allow us to understand the link between '*computer crash*' and '*rebooting my system'*. See Applicant's disclosure at page 55, ll. 20 – 25.

In other words, in the Applicant's disclosure, "semantic decoding" refers to a process which allows a natural language engine to compare two word sentences - *which may in fact have no words in common* - and yet still determine the meaning of the

user's utterance. This is done by semantically "decoding" at least some of the words in the utterance.[2]

Thus it can be seen that the "co-occurrence" filter in Kupiec et al. is not in fact "semantically decoding" any words in the user utterance. Instead, Kupiec et al is simply mechanically looking for whatever words in the utterance happen to be close to each other and then classifying them as "semantically" related. Unless this relationship is mirrored in the corpus documents, Kupiec et al. will not find a match for what the user said. The difference can be seen to be quite dramatic as Kupiec clearly cannot achieve the aims of the present claims. In the example given above from the Applicant's specification for instance, Kupiec would find no answer because it does not look at words in the query to decode such semantically to find semantic variants.[3]

The Junqua et al. reference is similarly deficient on this issue. While the reference explains that a global parser is used to interpret the meaning of the user's utterance, it notably fails to suggest that there is any kind of semantic decoding.

Accordingly the Examiner's combination of Kupiec et al. and Junqua et al. does not in fact even yield a combination that achieves the limitations of claim 1. One of the key criteria needed to establish *prima facie* obviousness of a claimed invention is that all the claim limitations must be taught or suggested by the prior art. *In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). "All words in a claim must be considered in judging the patentability of that claim against the prior art." *In re Wilson*, 424 F.2d 1382, 1385, 165 USPQ 494, 496 (CCPA 1970). *See, e.g.*, MPEP 2143.03; 706.02(j).

In this instance, the references clearly fail to disclose all limitations of the claim. On this basis, Applicant submits that the rejection cannot be sustained.

[2] It may not be necessary to semantically decode some keywords, or some noise words.

[3] Kupiec et al. does mention using a "thesaurus" and a table of "related" words to help disambiguate errors in transcriptions, as seen in col. 23, ll. 33 – 51. In the context described, however, it seems clear that he is referring to the use of the same for purposes of determining other words which look sound like the word that needs to be identified. For example, the phones re-co-n-ize b ea ch might be mapped to "recognize speech" or "wreck a nice beach." It can be seen that it is not semantically decoding these words, but merely finding homonyms, as he emphasizes that he is looking for "phonetically" represented/related words.

-- <u>It would not have been obvious to add “semantic decoding” and “statistical processing” from Kupiec’s speech recognition engine to the Junqua et al. natural language engine</u>

Even if the PTO disagrees that <u>Kupiec</u> does not teach “semantic decoding,” there is still a fundamental problem with adding such features unilaterally to the <u>Junqua et al</u>. <u>natural language</u> engine. To wit, the “semantic” aspects of <u>Kupiec’s</u> SR engine have been discussed. Nothing in <u>Kupiec</u> or <u>Junqua et al</u>. suggests how or why the NL engine in the latter could/would be modified to incorporate the type of co-occurrence filtering described in <u>Kupiec</u>. The Examiner does not provide any explanation of how/where such functionality could be integrated into J<u>unqua et al.</u>, let alone why such would then be incorporated by a skilled artisan into <u>Kupiec et al</u>. Consequently, other than the Examiner’s barebones conclusion that such would be useful for the “…purpose of permitting a user to communicate in any style of language” there is precious little guidance to one skilled in the art on how to achieve such result.

Moreover <u>Kupiec’s</u> “statistical processing” teachings identified by the Examiner are even more inapplicable to the type of natural language engine show in <u>Junqua et al</u>. In particular, <u>Kupiec</u> mentions that a “statistical model of transcriber errors…” could be developed, again, to increase transcription accuracy. See col. 23, ll. 46-51 and col. 9, ll. 52+:

> …hypothesis generator 60 can attempt to correct mistakes commonly made by transcriber 50 by adding, deleting or substituting one or more phones into the sequence of phones…such “correction” can be based, for example, on a **statistical model of the performance of transcriber 50**…”

Kupiec’s reference to <u>statistical processing</u>, therefore, is clearly limited to “phone” related corrections for transcribing <u>words</u>. There is simply no analogy given in either reference for a “natural language engine” “transcriber” that is error prone and would be similarly amenable to statistical modeling of the type noted in <u>Kupiec</u>. Since NL engines do not work on phones, what elements would it use instead? The references are simply too thin to support a rejection of this type. In this respect, therefore, the Examiner has again simply borrowed bits and pieces of the references to construct a combination that is not logically or technically consistent with any teachings of the disclosures.

Given the lack of any evidence which can fill in these missing gaps, the Applicant submits that the present record is insufficient to sustain an obviousness rejection.

A related but important fact is that Kupiec notes that while it can be used with either a discrete recognition system or a continuous speech recognition system (see col. 4, ll. 28 – 35) it clearly indicates that the embodiments therein were specifically adapted for a discrete speech recognition system (see col. 9, ll. 14+ discussing first embodiment; col. 18, ll. 35 – 37 discussing second embodiment) where discrete timing gaps must be provided by the user when speaking words. In fact Kupiec acknowledges the limitations of his approach when used in continuous speech recognitions systems (where users are free to speak continuously), by pointing out the increased computational load associated with compiling too many hypotheses of the user's utterance. See e.g., col. 21, l. 61 to col. 22, l. 20.

This is important because the Junqua et al. reference clearly specifies that the natural language engine therein is used in conjunction with a continuous speech recognition engine. See col. 3, ll. 46 – 55. Thus, as a first observation the Kupiec reference would seem to discourage one skilled in the art from using the kind of NL engine used in Junqua et al.

But more importantly the admission by Kupiec is key because it shows that one skilled in the art would be led away from trying to incorporate the additional computational load of a natural language engine (with the attendant new semantic and statistical processing the Examiner asserts it would need) on top of the computational load restrictions associated with the speech recognitions engine used in that reference. The Examiner effectively requires that the processing of Kupiec be multiplied several times without any rhyme or reason, thus directly impairing the performance of such system in a manner inconsistent with the teachings of such disclosure.

For these additional reasons Applicant submits that the references are not properly combinable because the teachings of Kupiec et al. are not extendible to Junqua et al. Moreover the references appear to teach away from such combination, which is the hallmark of non-obviousness.

-- It would not have been obvious to add "search predicates" from Kupiec's speech recognition engine to the Junqua et al. natural language engine

The Examiner has argued that it would have been obvious to incorporate the "search predicates" generated by Kupiec's speech engine into a natural language engine of the type described in Junqua et al. This again is a simplistic analysis of the references and ignores entirely the differences in these two different types of technologies as well as the language in the claims.

First, as noted above, Kupiec uses co-occurrence data to identify words spoken in an utterance. To determine this "co-occurrence" he uses certain *proximity* operators as noted in col. 11, ll. 9 – 41 as well as certain Boolean operators to find text matches in the corpus.

There is simply nothing in either Kupiec or Junqua et al. to suggest to one skilled in the art that they should modify a natural language engine to include this type of capability. Nor is there anything in the Office Actions which explains or documents with substantive evidence why such combination would be desirable.

Finally, the Applicant further notes that the Examiner has not given proper weight to this limitation of the claim, which recites

> "…a natural language engine configured for linguistically processing *said recognized words* to *generate search predicates* for said articulated speech utterance;"

As set out in this limitation therefore, the claim makes clear that the search predicates are derived from recognized words, and not simply hypotheses of what these words might be as is done in Kupiec et al. Again, since Kupiec is directed to speech recognition only, it is apparent that his search predicates are being used to actually identify the words, which means that they are not yet recognized when the predicates are generated. In contrast the present claims make clear that the search predicates are derived from another processing step after words are actually recognized.

Consequently Applicants submit that this provides yet another reason why claim 1 cannot be made obvious by this combination of references.

-- It would not have been obvious to add any linguistic processing from Kupiec into Junqua et al. for processing a ".. set of one or more corresponding recognized matches to determine a final match"

As an initial matter it should be noted that any linguistic processing in Kupiec is apparently quite limited as the Examiner is citing to the co-occurrence matching in the document corpus for recognizing words. The object of the Junqua et al. natural language processor is to interpret the meaning of the user's utterance, and it does this by mapping words to tags/tasks. As the PTO can determine from reading such reference, the usefulness or benefit of using a word co-occurrence function for this type is extremely unclear.

To this end in fact the Examiner does not cite to any evidence on how or why it would be obvious to include such type of co-occurrence operation in the Junqua et al. natural language engine. It can be seen, quite clearly in fact that the latter does not linguistically process the potential *interpretations* of the user's utterance. Yet the Examiner, without any explanation or support, concludes that one skilled in the art would simply modify the Junqua et al. natural language engine to impart a capability from Kupiec that is not called for.

Again the Applicant submits that this is yet another instance of the PTO combining two references which are inherently incompatible because they are directed to different problems (speech recognition and natural language processing). Under the circumstances Applicant suggests that an obviousness rejection is not supportable for this claim or the claims dependent therefrom (2 – 14).

Dependent claims 2 - 14

With respect to the dependent claims, Applicant submits the following additional reasons for patentability:

For claim 2, Applicant submits that Kupiec et al. does not show any first level query with words alone, then followed by a second level query with predicates. By the Examiner's logic, Kupiec only shows queries which involve predicates generated by the search engine.

For claim 3 nothing in Kupiec suggests this type of operation. If the Examiner contends that the queries in Kupiec include "search predicates" then it is apparent that the latter must be created before the queries are executed, and cannot be done at the same time.

For claim 4: (Kupiec (5,500,920) taken with Junqua et al. (6,314,398) taken with

McDonough et al. (5,625,748)): McDonough et al. does not cure any of the deficiencies of the other references in this regard. McDonough et al. is apparently generally discussing a mapping of messages to topics. The Examiner states that McDonough et al. employs a Kullback-Liebler distance measure; but there is no indication that this is the same as claim 4 in which a specific term frequency calculation is based on calculating a lexical distance between each word of the recognized query and the topic query entries.

Dependent claim 12: (Kupiec (5,500,920) taken with Junqua et al. (6,314,398) and Joost) this claim makes explicit that the data is optimized on a query by query basis. Joost is not believed to cure the deficiencies of the other references in this regard, and does not appear to show any degree of optimization as set out in claim 12.

Independent claim 22

While independent claim 22 is a method claim and does not recite the "natural language engine" structural limitations of claim 1, the arguments are substantially the same because the main language in dispute is the same in both cases. For example claim 22 also requires that the linguistic processing be performed on recognized text for the search predicates; and that both "semantic decoding" and statistical processing are used to determine a final match.

In addition claim 22 includes a further operation specifying that there is a comparison between potential matches and the speech utterance. In Kupiec et al. there is little or no explanation of what is done after the query locates the appropriate documents, let alone that there is a linguistic processing done on such documents for a comparison. What the reference says in fact is merely that the documents retrieved in the text corpus search "…can also be analyzed and evaluated for relevance to the preferred hypothesis or hypotheses in applications where this is appropriate." See col. 23, ll. 12 -15. Thus, there is no description of what analysis might be used.

Accordingly the Applicant respectfully submits that this is yet another reason why this claim and its dependents (23 – 28) would not have been obvious.

Dependent claims 23 – 28

Claim 24: see claim 3.

Claim 25: see claim 2.

Claim 28: Applicant's reading of Joost suggests that it does not disclose that multiple servers perform the same operations; only that a particular server is selected to perform the required operations.

IX. CLAIMS

A copy of the claims involved in the present appeal is attached hereto as Appendix A.

X. EVIDENCE

No additional evidence pursuant to §§ 1.130, 1.131 or 1.132 or entered by or relied upon by the Examiner is being submitted.

XI. RELATED PROCEEDINGS

No related proceedings are referenced herein, nor are copies of decisions in related proceedings being provided, as there are none. Accordingly, no Appendix is included.

Respectfully submitted,

J. Nicholas Gross
Registration No. 34,175
Attorney for Applicant(s)

April 7, 2008
2030 Addison Street
Suite 610
Berkeley, CA 94704
510-540-6300
510-540-6315 (fax)

APPENDIX A

1. (Previously amended) A speech query recognition system comprising:

   a speech recognition engine for generating recognized words taken from an articulated speech utterance; and

   a natural language engine configured for linguistically processing said recognized words to generate at least two different types of search predicates for said articulated speech utterance;

   wherein said search predicates correspond to logical operators to be satisfied by a potential recognition match;

   a query formulation engine adapted to convert said recognized words and said search predicates into a structured query suitable for locating a set of one or more corresponding recognized matches for said articulated speech utterance; and

   said natural language engine further being configured for linguistically processing said set of one or more corresponding recognized matches to determine a final match for said articulated speech utterance using both semantic decoding and statistical based processing performed on said recognized words.

2. (Original) The system of claim 1, wherein said query formulation engine generates a first level query using said recognized words alone, and further customizes said first level query using said search predicates to generate a second level query.

3. (Original) The system of claim 2, wherein said natural language operates to generate said search predicates during a time when said query formulation engine generates said first level query.

4. (Original) The system of claim 1, wherein said semantic decoding is based on a term frequency calculation, which term frequency calculation is based on calculating a lexical distance between each word in said recognized words with each word of one or more topic query entries.

5. (Original) The system of claim 1, wherein said natural language engines uses a set of context parameters for generating said search predicates.

6. (Original) The system of claim 1, wherein said speech recognition engine, said natural language engine, and said query formulation engine are implemented as routines executing on a server computing system.
7. (Original) The system of claim 1, wherein said statistical based processing includes an operation for determining noun-phrases in said speech utterance.
8. (Original) The system of claim 1, wherein said natural language engine compares noun-phrases of said set of potential matches with noun-phrases of said speech utterance to determine said final match.
9. (Original) The system of claim 1, wherein said final match is determined in real-time.
10. (Original) The system of claim 9, wherein said speech utterance can correspond to one of more than 100 potential corresponding potential matches, and said final match is used for determining an articulated answer to said speech utterance in less than 10 seconds.
11. (Original) The system of claim 1, wherein said speech recognition is distributed across a client-server architecture.
12. (Previously amended) The system of claim 11, wherein said client generates an amount of speech data that is optimized on a query by query basis to reduce recognition latencies.
13. (Original) The system of claim 1, wherein said recognized speech utterance is used for controlling a web page.
14. (Original) The system of claim 1, wherein said structured query is a full text query containing SQL search predicates.
15. (Original) The system of claim 1, wherein said corresponding potential matches are retrieved from a relational database that is updated asynchronously to reduce retrieval latency.

16 – 21 (Canceled)

22. (Previously amended) A method of recognizing a speech query comprising the steps of:

(a) recognizing text in an articulated speech utterance; and

(b) linguistically processing said recognized text to generate at least two different types of search predicates for said articulate speech utterance;

wherein said search predicates correspond to logical operators to be satisfied by a potential recognition match;

(c) generating a query to identify a potential match for said speech utterance, said query being based on said recognized text and said search predicates;

(d) determining a final match for said speech utterance by linguistically comparing any potential matches identified by said query with said articulated speech utterance;

wherein both semantic decoding and statistical based processing operations are used to determine said final match.

23. (Original) The method of claim 22, further including a step: (e) retrieving a matching response for said final match, which matching response is provided in audible form.

24. (Original) The method of claim 22, wherein steps (b) and (c) overlap in time.

25. (Original) The method of claim 24, wherein step (c) includes two sub-steps, including a step (c)' wherein a preliminary query is generated based on said recognized text, and a step (c)'' wherein a final query is generated based on said preliminary query and said search predicates.

26. (Original) The method of claim 22, wherein said final match is determined by comparing noun-phrases of said speech utterance and said potential matches.

27. (Original) The method of claim 22, wherein step (a) occurs across a distributed computing platform, including a client device and a server device.

28. (Original) The method of claim 22, wherein steps (a) to (d) occur simultaneously across multiple servers in response to a speech utterance from a single client device.